

SIMULATING PANDEMIC INFLUENZA RISKS OF US CITIES

Catherine Dibble
Stephen Wendel
Kristofor Carle

Department of Geography
2181 LeFrak Hall
University of Maryland
College Park, M.D. 20742 U.S.A.

ABSTRACT

We describe the spatial Agent-Based Computational Laboratory that we have developed to study the pandemic influenza risks of US cities. This research presented a series of interesting challenges, from the uncertainty surrounding the future epidemiological characteristics of a human-transmission H5N1 strain of pandemic influenza, to the need to provide timely decision-support despite modeling a country with a population of 300 million individuals. In order to provide time-sensitive policy analyses, we implemented a light-and-fast agent-based model of the spatial and temporal spread of pandemic influenza, which uses a novel compression technique to analyze large numbers of agents. We assessed the impact of parameter uncertainty and of stochastic behavior via very large numbers of simulations. To facilitate this, we developed a parallel job controller that tests combinations of disease scenarios, and a platform-independent job-submission application that harnesses the computational resources of high-performance computing environments ranging from local clusters up through TeraGrid super-computers.

1 INTRODUCTION

H5N1 influenza has killed 61% (189 of 310) of its human victims (WHO 2007). Although it does not yet spread easily from human to human, mutation and viral reassortment among flu strains is likely to generate a human-adapted variant capable of causing a world-wide influenza pandemic more severe than the pandemic of 1918-1919. The National Academies Institute of Medicine convened an emergency panel of experts to produce a letter report evaluating the usefulness of detailed community-level simulation models for effective community containment of pandemic influenza (Institute of Medicine 2006). Their report emphasizes the usefulness and desir-

ability of complementary simulations with simpler models capable of incorporating surveillance information and fast enough to provide policy-relevant updates and risk evaluations during a pandemic. This extended abstract describes the light-and-fast agent-based simulation model and spatial agent-based Computational Laboratory we developed to help prepare for a pandemic.

2 BACKGROUND

There is a rich history of work on the spread of disease, including considerable recent work with agent-based simulation models. The widely used SEIR model provides a set of differential equations that approximate disease spread (Anderson and May 1991). The model predicts the mean quantity of four categories of individuals—Susceptible, Exposed, Immune, and Recovered—who interact over time within a single perfectly mixed population. Keeling and Grenfell (2000) examined the role of discretization in SEIR models, finding that discretization provides a more accurate approximation of disease spread. Authors such as Lloyd and May (1996) have examined the impact of spatial heterogeneity and stochasticity on the spread of the disease, similarly verifying the importance of these extensions.

In search of improved predictive power, a number of recent projects have developed fine-grain computational models where each individual is modeled separately, and interact with peers according to local contact processes. Notable work includes Germann et al. (2006) and Ferguson et al. (2006), both of which present models of the entire United States calibrated to community level demographics and behavior. These models provide a previously impossible level of detail, yet they have two major drawbacks. First, their computational requirements are immense. Second, and related to the first, their considerable complexity makes it difficult to study the sensitivity of the model to its assumptions and input parameters. This extended abstract presents an alternative,

complementary, approach to community-level models, which focuses on a light-and-fast agent-based model embedded in a computational laboratory to generate timely policy feedback.

3 RESEARCH DESIGN

3.1 Agent-Based Model

Modeling the spread of pandemic influenza presents unusual challenges, from the uncertainty surrounding the future epidemiological characteristics of a human-transmission H5N1 strain, to the need to model a country with a population of 300 million individuals. We developed a compact, efficient agent-based model of the geographic spread of influenza within the United States, which we then placed in a computational laboratory to evaluate the impact of the uncertainty resulting from epidemiological characteristics and stochastic travel and transmission behavior.

The epidemiological component of the model consists of a discretized, spatialized, agent-based SEIR simulation model, where agents follow carefully calibrated travel behavior and each has a set of internal clocks to govern the progression of disease phases from susceptible through exposed, infectious, and recovered or removed. The agents are proportionally allocated to the fifty most populous metropolitan areas of the United States. Agent travel among the metropolitan areas is carefully calibrated according to a representative sample of US air travel behavior. Within each metropolitan area, the agents interact randomly with one another. The core model thus builds upon previous work on stochastic, discrete, spatial SEIR models, but adds the capacity for agent heterogeneity and customized travel behavior.

The fifty metropolitan areas included in the model represent roughly 50% of the US population, or 150 million agents. We developed a novel compression technique, Dynamic Agent Compression, to represent each individual as a potentially heterogeneous, infected agent in the model (Wendel and Dibble 2007). We testing the sensitivity of the model's results to the number of agents from 150 million down to 100, and found that with appropriate scaling factors a population as small as 100,000 agents provided unbiased results.

3.2 Computational Laboratory

We put the model through its paces using a Computational Laboratory to control and evaluate hundreds of thousands of simulations in order to assess the scaling sensitivity for numbers of agents, to evaluate the effects of parameter uncertainty and of stochastic behavior, to search for best-case and worst-case scenarios, and to op-

timize effective geographic deployments. The Computational Laboratory consists of calibration utilities to process empirical input data, an automated controller to execute the model, a supervisory genetic algorithm for search and optimization, and a set of statistical tools to analyze simulation results.

The controller performed sweeps over specific ranges of parameters, to test the local sensitivity of the model. Since a single execution of a stochastic model is nearly meaningless, the controller automatically executed the model over a range of random seeds for each unique set of input parameters. An associated job-submission application provided a platform-independent interface to harness the computational resources of high-performance computing environments from local clusters up through TeraGrid super-computers (TeraGrid 2007).

We employed a supervisory genetic algorithm to optimize geographic deployments of pandemic interventions and to evaluate best-case and worst-case stochastic outcomes. This allowed us to determine optimal allocations of resources with far greater efficiency than a brute force combinatorial parameter sweep. The genetic algorithm can be set to search across combinations of key parameters for extreme values of single or multiple combinations of outcome variables, based on results from one or more stochastic replications. The greater economy of this supervisory search and optimization tool releases computational resources that may in turn be used to simulate sufficient stochastic replications to be able to distinguish statistically significant differences among scenario outcomes (Dibble 2006).

4 SUMMARY

We have found that our Computational Laboratory can effectively explore the sensitivity of a model to its parameters and stochastic events. The agent-based model used to simulate pandemic influenza in this project is efficient enough to provide rapid results across the many thousands of replications and parameter sets employed in the study. During the presentation, we will discuss some of the key lessons learned with this computational laboratory for pandemic influenza.

ACKNOWLEDGMENTS

We would like to thank the Department of Health and Human Services National Vaccine Program Office, the Office of Naval Research, Grant N000140310062, the National Institutes of Health, National Institute of General Medical Sciences, Models of Infectious Disease Agent Study (MIDAS), Grant 1 U01 GM070698, for supporting this work. We would also like to thank the TeraGrid Development Allocation Committee (DAC), the University of Maryland Office of Information Technology

(OIT), and the University of Maryland Lattice Project for providing computational resources.

REFERENCES

- Anderson, R. M., and R. M. May. 1991. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press.
- Dibble, C. 2006. Computational Laboratories for Spatial Agent-Based Models. *Handbook of Computational Economics II: Agent-Based Computational Economics*, Leigh Tesfatsion and Kenneth L. Judd, Editors, North-Holland.
- Ferguson, N. M., D. A. Cummings, C. Fraser, J. C. Cajka, P. C. Cooley, and D. S. Burke. 2006. Strategies for mitigating an influenza pandemic. *Nature* 442(7101): 448-52.
- Germann, T. C., K. Kadau, I. M. Longini, and C. A. Macken. 2006. Mitigation strategies for pandemic influenza in the United States. *PNAS* 103(15): 5935-40.
- Institute of Medicine, Committee on Modeling Community Containment for Pandemic Influenza. 2006. *Modeling Community Containment for Pandemic Influenza: A Letter Report*, The National Academies Press, Washington D.C.
- Keeling, M. J., and B. T. Grenfell. 2000. Individual-based Perspectives on R0. *Journal of Theoretical Biology*, no. 2003: 51-61.
- Lloyd, A. L., and R. M. May. 1996. Spatial Heterogeneity in Epidemic Models. *Journal of Theoretical Biology* 179: 1-11.
- The TeraGrid Project 2007. TeraGrid Home Page Available via <http://www.teragrid.org/> [accessed July 2007].
- Wendel, S. and C. Dibble. 2007. Dynamic Agent Compression. *Journal of Artificial Societies and Social Simulation* 10(2):9 <http://jasss.soc.surrey.ac.uk/10/2/9.html>
- World Health Organization. 2007. "Cumulative Number of Confirmed Human Cases of Avian Influenza A/(H5N1) Reported to WHO." Available via http://www.who.int/csr/disease/avian_influenza/country/cases_table_2007_06_06/en/index.html [accessed 6 June 2007].

AUTHOR BIOGRAPHIES

CATHERINE DIBBLE is an Assistant Professor in the Department of Geography at the University of Maryland College Park. She has been working professionally with spatial evolutionary algorithms, location-allocation optimization, and adaptive spatio-temporal relevance filters since the early 1990s, and with spatial agent-based computational laboratories since the late 1990s. She serves on an NRC Committee for Organizational Modeling from Individuals to Societies, and is a co-Investigator with the

NIH-funded MIDAS project for controlling epidemics. Dr. Dibble teaches courses in Economic Geography and Computational Laboratories, and leads an active Computational Laboratories research group at the University of Maryland.

STEPHEN WENDEL is the research manager for Catherine Dibble's Computational Laboratories research team at the University of Maryland, College Park. His background includes software engineering at Pixar Animation Studios, field work in South Africa on the application of technology for economic development, and research into economic integration. He holds an MA from Johns Hopkins-SAIS and is currently pursuing a PhD at the University of Maryland's Department of Government and Politics, where he uses agent-based modeling to study political radicalization.

KRIS CARLE is a Faculty Research Assistant with Catherine Dibble's Computational Laboratories research team. He attended the University of Maryland and completed a Bachelor's degree with a double major in computer science and geography, specializing in GIS. During college Kris worked as an operations manager and transit planner for Shuttle-UM transit system at the University of Maryland. His academic interests include transportation operations and networks, and computer networks and security. He hopes to continue his education and earn a PhD.